

# On Moments of Discrete Phase-Type Distributions

Tuğrul Dayar

Department of Computer Engineering, Bilkent University,  
TR-06800 Bilkent, Ankara, Turkey  
`tugrul@cs.bilkent.edu.tr`

**Abstract.** Recently, an efficient and stable method to compute moments of first passage times from a subset of states classified as safe to the other states in ergodic discrete-time Markov chains (DTMCs) has been proposed. This paper shows that the same method can be used to compute moments of discrete phase-type (DPH) distributions, analyzes its complexity on various acyclic DPH (ADPH) distributions, and presents results on a set of DPH distributions arising in a test suite of DTMCs.

## 1 Introduction

Discrete and continuous phase-type (PH) distributions have been introduced by Neuts [12,13] more than two decades ago. Although continuous PH distributions have been extensively used over the years in stochastic modeling, discrete PH (DPH) distributions have not been as widely studied, and there seems to be room for research to understand their merits and properties (see, for instance, [1,2,16]). This paper aims at rectifying the situation from a moment computation point of view.

Recently, an efficient and stable method to compute moments of first passage times from a subset of states classified as safe to the other states in ergodic discrete-time Markov chains (DTMCs) has been proposed in [4]. Although a recurrence for the moment vector of first passage times involving a common coefficient matrix and the previous moment vector on the right-hand side for the continuous-time case had been known for some time (see, for instance, [9]), a similar recurrence had not been given for the discrete-time case. The difficulty in the discrete-time case lied in the fact that equations involved factorial moments rather than moments and there was no obvious way to go from factorial moments to moments. The method proposed in [4] reduced the computation of the first  $k$  moment vectors of first passage times from a subset of states to the solution of  $k$  linear systems with a common coefficient matrix and right-hand sides governed by a novel recurrence involving the binomial coefficients and the first  $(k-1)$  moment vectors. The efficiency and stability of the proposed method rest respectively on the smaller linear systems solved with a common coefficient matrix and the particular implementation of the Grassmann-Taksar-Heyman (GTH) method [8] used in factorizing the coefficient matrix.

Since the first passage time from a given set of states to the other states in an ergodic DTMC is also the absorption time in the other states, it is possible to use the method proposed in [4] to compute moments of DPH distributions. It is our belief that moment matching techniques will benefit from the ability to compute higher moments of DPH distributions without having to resort to generating functions, transform techniques, or factorial moments. Therefore, in this paper we investigate the effects of using the proposed method of computing moments on various DPH distributions by concentrating on complexity issues.

The organization of the paper is as follows. In section 2, we provide a background on DPH distributions and introduce the moment computation method. In section 3, we give two examples of nonacyclic DPH distributions that have been used in stochastic modeling. In section 4, we investigate the computation of moments of acyclic DPH distributions and give an example. In section 5, we present results on moments of DPH distributions obtained from a test suite of DTMCs, and in section 6 we conclude.

## 2 Discrete Phase-Type Distributions

Let  $H$  be the random variable associated with the time to absorption in a finite DTMC,  $P$ , of order  $(n_S + 1)$  with  $n_S$  transient states and one absorbing state conditioned on an initial probability (row) vector,  $\tau$ . Then  $f_H(h) = \Pr\{H = h\}$  is said to be the probability mass function of a discrete phase-type (DPH) distribution of order  $n_S$ .

Let us assume that the transient states are those in  $\mathcal{S} = \{0, 1, \dots, n_S - 1\}$  and the absorbing state is  $n_S$ ; otherwise, the states can always be renumbered. In matrix form, we have

$$P = \begin{pmatrix} P_{\mathcal{S},\mathcal{S}} & p_{\mathcal{S}} \\ 0^T & 1 \end{pmatrix}, \quad (1)$$

where  $p_{\mathcal{S}} = (I - P_{\mathcal{S},\mathcal{S}})e$  and  $e$  is the column vector of ones of appropriate length. Now, let  $\tau_{\mathcal{S}} = (\tau_0, \tau_1, \dots, \tau_{n_S-1})$ , and note that  $P_{\mathcal{S},\mathcal{S}}$  and  $\tau_{\mathcal{S}}$  uniquely determine  $p_{\mathcal{S}}$  and  $\tau_{n_S}$ , respectively. The pair  $(\tau_{\mathcal{S}}, P_{\mathcal{S},\mathcal{S}})$  is called the representation of the DPH random variable  $H$ . When  $P_{\mathcal{S},\mathcal{S}}$  (can be symmetrically permuted to or) is in upper-triangular form, the DPH distribution is said to be acyclic [1].

In this paper, for brevity, we consider only those DPH distributions for which  $\tau_{n_S} = 0$  (implying that an absorption time of zero in (1) is not possible), and therefore we have

$$f_H(h) = \Pr\{H = h\} = \tau_{\mathcal{S}} P_{\mathcal{S},\mathcal{S}}^{h-1} p_{\mathcal{S}} \quad \text{for } h \geq 1. \quad (2)$$

When a DPH distribution  $(\tau_{\mathcal{S}}, P_{\mathcal{S},\mathcal{S}})$  is used to model a stochastic process which upon absorption starts anew according to the initial probability distribution  $\tau_{\mathcal{S}}$ , we say we have a DPH renewal process with interrenewal times that are DPH distributed.

Noticing that  $H$  with the probability mass function in (2) is also the random variable associated with time of first passage from the transient states in  $\mathcal{S}$  to the absorbing state  $n_{\mathcal{S}}$  as in [4], its  $(i+1)$ st moment is given by

$$E[H^{i+1}] = \sum_{h=0}^{\infty} h^{i+1} f_H(h) = \tau_{\mathcal{S}} m^{(i+1)} \quad \text{for } i \geq 0. \quad (3)$$

Here,  $m^{(i+1)}$  is the  $(i+1)$ st moment vector of time to absorption and satisfies the recurrence

$$(I - P_{\mathcal{S},\mathcal{S}})m^{(i+1)} = \sum_{j=0}^i (-1)^{i-j} \binom{i+1}{j} m^{(j)} \quad \text{for } i \geq 0 \quad \text{with } m^{(0)} = e \quad (4)$$

in which

$$\binom{i+1}{j} = \frac{(i+1)!}{j!(i+1-j)!}$$

denotes the binomial coefficient in Pascal's triangle that is read “ $(i+1)$  choose  $j$ ” [7]. The proof of this result appears in the appendix of [4] and follows from various identities involving binomial coefficients and Stirling numbers [7]. In passing, we remark that the coefficient matrix,  $(I - P_{\mathcal{S},\mathcal{S}})$ , of the linear system in (4) is nonsingular when  $P$  is ergodic (due to the irreducibility assumption in its definition).

Using the identity

$$\binom{i+1}{j} = \binom{i}{j-1} + \binom{i}{j}$$

of binomial coefficients [7], the following algorithm to compute the moment vectors,  $m^{(i+1)}$ ,  $i \geq 0$ , has been given in [4]. Here, we have added one more line to compute the moments,  $E[H^{i+1}]$ ,  $i \geq 0$ , as well. When a DPH distribution is used to model a stochastic process which upon absorption starts anew according to the given initial probability distribution, Algorithm 1 computes moment vectors and moments associated with DPH distributed interrenewal times.

#### ALGORITHM 1

*Computing the  $k$ th moment vector and moment of a DPH distribution.*

```

LU factorize  $(I - P_{\mathcal{S},\mathcal{S}})$  using GTH;
 $a_0 = 1$ ;  $a_1 = -1$ ;  $m^{(0)} = e$ ;
Solve for  $m^{(1)}$  in  $LUm^{(1)} = m^{(0)}$ ;
For  $i = 1$  up to  $k-1$  {
     $a_{i+1} = -1$ ;
    For  $j = i$  down to 1
         $a_j = a_{j-1} - a_j$ ;
     $a_0 = -a_0$ ;
    Solve for  $m^{(i+1)}$  in  $LUm^{(i+1)} = \sum_{j=0}^i a_j m^{(j)}$ ;
     $E[H^{i+1}] = \tau_{\mathcal{S}} m^{(i+1)}$ ;
}
```

The order,  $n_S$ , of the DPH distribution is mostly small, in the order of tens, thereby suggesting the use of a direct method for the solution of (4). The GTH implementation proposed in [4] for the factorization  $(I - P_{S,S}) = LU$  at the outset of Algorithm 1 executes on a nonhomogeneous, but consistent linear system with a singular coefficient matrix of order  $(n_S + 1)$  in the form of a modified Gaussian elimination. Other than ensuring stability, this implementation turns out to be very efficient since it can be coded completely in row-wise sparse format with delayed row updates [15].

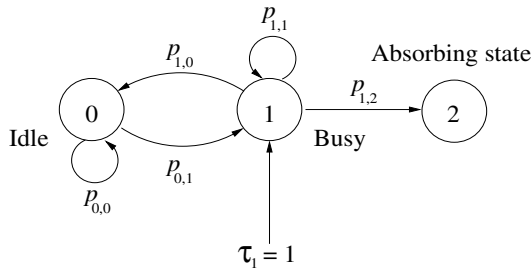
Now, let  $nz_{LU}$  denote the sum of the nonzeros in the strictly lower-triangular part of  $L$  (since ones along its diagonal are not stored) and  $U$ . Also let  $k$  denote the order of the highest moment vector to be computed. Then, in addition to the time to sparse factorize  $(I - P_{S,S})$  at the outset, which is cubic in  $n_S$  in the worst case when the matrix is full, Algorithm 1 takes  $(k - 1)[2nz_{LU} + n_S(k + 3) + k/2]$  floating-point arithmetic operations for a total of  $(k - 1)$  passes. In terms of space complexity, it requires  $(nz_{LU} + (k + 2)n_S + (k + 1))$  floating-point and  $(nz_{LU} + n_S + 1)$  integer storage space. See [4] for details. Here, we have also accounted for the time and space required by the computation of each moment as in the last line of the outer for-loop of Algorithm 1.

In the next section, we give two examples of nonacyclic DPH distributions used as renewal processes.

### 3 Examples

We provide two DPH distributions that have been used in traffic modeling.

*Example 1. The interrupted Bernoulli process (IBP).* When the arrival of pack-



**Fig. 1.** The IBP interrenewal distribution

ets to a node in a communication network is of bursty nature, one may use the interrupted Bernoulli process (IBP) to model the arrival of packets. This form of arrivals is inspired by the fact that, most of the time packets arrive to a node from several external sources and they happen to be grouped in bursts.

The IBP is a DPH renewal process which is governed by busy and idle periods. Time is thought of as being slotted and state changes may only occur at points

that are multiples of a slot duration. The IBP interrenewal distribution in Figure 1 is represented by the DTMC

$$P_{IBP} = \left( \begin{array}{cc|c} p_{0,0} & p_{0,1} & 0 \\ \hline p_{1,0} & p_{1,1} & p_{1,2} \\ 0 & 0 & 1 \end{array} \right), \quad (5)$$

whose  $(2 \times 2)$  principal submatrix is nonacyclic, and the initial probability vector  $\tau_S = (0, 1)$ . After an absorption, the DPH renewal process starts anew so as to remain in the busy period where renewals (that is, arrivals) can take place.

It is also possible to view the transition probabilities in (5) as given by

$$p_{0,1} = \beta, \quad p_{1,0} = \alpha, \quad \text{and} \quad p_{1,2} = (1 - \alpha)\lambda,$$

where  $\lambda$  is the probability of a renewal in the particular slot within a busy period [11]. Note that when  $\lambda = 1$ , the self-loop in state 1 does not exist. This is understandable since then at each slot in the busy period a renewal takes place.

Using the  $(2 \times 2)$  principal submatrix of  $P_{IBP}$ , we obtain

$$I - P_{S,S} = \begin{pmatrix} 1 - p_{0,0} & -p_{0,1} \\ -p_{1,0} & 1 - p_{1,1} \end{pmatrix}$$

which has the  $L$  and  $U$  factors

$$L = \begin{pmatrix} 1 & 0 \\ -p_{1,0}/(1 - p_{0,0}) & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 - p_{0,0} & -p_{0,1} \\ 0 & 1 - p_{1,1} + p_{0,1}p_{1,0}/(1 - p_{0,0}) \end{pmatrix}.$$

Then the remaining steps of Algorithm 1 can be executed to compute moments of the IBP interrenewal distribution.

*Example 2. Markov modulated Bernoulli process (MMBP).* Consider a slightly different version of the Markov modulated Bernoulli process (MMBP) that has been proposed to emulate self-similarity in [14] with initial probability vector  $\tau_S = (0, 0, \dots, 0, 1)$  and absorbing DTMC

$$P_{MMBP} = \left( \begin{array}{cccc|c} \sum_{n_S-1} & 0 & \cdots & 0 & (q/a)^{n_S-1} & 0 \\ 0 & \sum_{n_S-2} & \cdots & 0 & (q/a)^{n_S-2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sum_1 & (q/a) & 0 \\ \hline 1/a^{n_S-1} & 1/a^{n_S-2} & \cdots & 1/a & (1-\lambda)\sum_0 & \lambda\sum_0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right), \quad (6)$$

where  $0 < q \leq a$ ,  $1 < a$ ,  $0 < \lambda \leq 1$ ,  $\sum_0 = 1 - \sum_{j=0}^{n_S-1} 1/a^j$  (hence,  $0 < \sum_0 < 1$ ), and  $\sum_i = 1 - (q/a)^i$  for  $1 \leq i \leq n_S - 1$ . Observe that IBP is a special case of MMBP in which  $n_S = 2$ ,  $\alpha = 1/a$ , and  $\beta = q/a$ .

Using the nonacyclic  $(n_S \times n_S)$  principal submatrix of  $P_{MMBP}$  in (6), we obtain

$$I - P_{S,S} = \begin{pmatrix} 1 - \sum_{n_S-1} & 0 & \cdots & 0 & -(q/a)^{n_S-1} \\ 0 & 1 - \sum_{n_S-2} & \cdots & 0 & -(q/a)^{n_S-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 - \sum_1 & -(q/a) \\ -1/a^{n_S-1} & -1/a^{n_S-2} & \cdots & -1/a & 1 - (1-\lambda) \sum_0 \end{pmatrix}$$

which has the  $L$  and  $U$  factors

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \frac{-1}{a^{n_S-1}(1-\sum_{n_S-1})} & \frac{-1}{a^{n_S-2}(1-\sum_{n_S-2})} & \cdots & \frac{-1}{a(1-\sum_1)} & 1 \end{pmatrix},$$

$$U = \begin{pmatrix} 1 - \sum_{n_S-1} & 0 & \cdots & 0 & u_{0,n_S-1} \\ 0 & 1 - \sum_{n_S-2} & \cdots & 0 & u_{1,n_S-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 - \sum_1 & u_{n_S-2,n_S-1} \\ 0 & 0 & \cdots & 0 & u_{n_S-1,n_S-1} \end{pmatrix},$$

where

$$u_{i,n_S-1} = -(q/a)^i - \sum_{j=0}^{i-1} u_{j,n_S-1} / (a^{n_S-j} (1 - \sum_{n_S-j})) \quad \text{for } 0 \leq i \leq n_S - 2,$$

$$u_{n_S-1,n_S-1} = 1 - (1-\lambda) \sum_0 - \sum_{j=0}^{n_S-2} u_{j,n_S-1} / (a^{n_S-j} (1 - \sum_{n_S-j})).$$

Then the remaining steps of Algorithm 1 can be executed to compute the moments of the MMBP interrenewal distribution.

See, for instance, section 2.5 of [10] for other examples, and observe that there is a lot of sparsity in  $P_{S,S}$  as in Example 2. In fact, no fill-in (i.e., nonzeros in the  $L$  and  $U$  factors replacing zeros in the coefficient matrix) is introduced during the LU factorization of  $(I - P_{S,S})$  since it is a matrix with arrowhead nonzero structure.

Although we were able to give explicitly the LU factorization of  $(I - P_{S,S})$  for the examples considered in this section, in general this cannot be done, and the factorization needs to be performed numerically using the GTH implementation discussed in [4]. However, for the class of DPH distributions discussed in the next section, the factorization is available explicitly.

## 4 Acyclic Discrete Phase-Type Distributions

We consider three types of acyclic discrete phase-type (ADPH) distributions. It is our assumption that the matrix  $P_{S,S}$  in the ADPH representation is already in upper-triangular form. This implies that there is no need to LU factorize  $(I - P_{S,S})$  at the outset of Algorithm 1. In other words,  $L = I$  and  $U = (I - P_{S,S})$ .

Now, let us define

$$b^{(i+1)} = \sum_{j=0}^i a_j m^{(j)} \quad \text{for } i \geq 0 \quad \text{with } m^{(0)} = e.$$

In the general acyclic case, when  $(I - P_{S,S})$  is upper-triangular, the last to next step in the outer for-loop of Algorithm 1 can be rewritten as

$$m_k^{(i+1)} = (b_k^{(i+1)} + \sum_{k'=k+1}^{n_S-1} p_{k,k'} m_{k'}^{(i+1)}) / (1 - p_{k,k}) \quad \text{for } k = n_S - 1 \text{ down to } 0.$$

When  $(I - P_{S,S})$  is upper-bidiagonal, the same step can be rewritten as

$$\begin{aligned} m_{n_S}^{(i+1)} &= b_{n_S-1}^{(i+1)} / (1 - p_{n_S-1, n_S-1}), \\ m_k^{(i+1)} &= (b_k^{(i+1)} + p_{k, k+1} m_{k+1}^{(i+1)}) / (1 - p_{k,k}) \quad \text{for } k = n_S - 2 \text{ down to } 0. \end{aligned}$$

Note that this is the case when the ADPH distribution is discrete Erlang (DE) with  $n_S$  phases. When  $S$  is a diagonal matrix, the last step reduces to

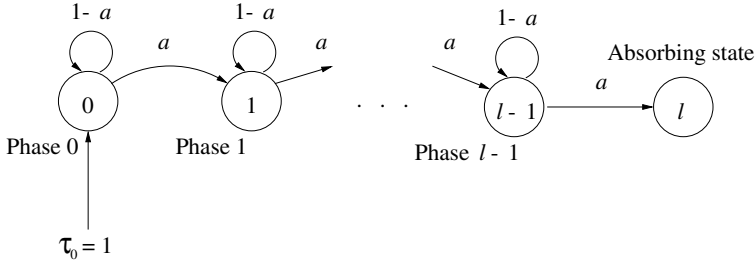
$$m_k^{(i+1)} = b_k^{(i+1)} / (1 - p_{k,k}) \quad \text{for } 0 \leq k \leq n_S - 1.$$

Then, for a total of  $(k-1)$  passes Algorithm 1 takes  $(k-1)[2nz_U + n_s(k+3) + k/2]$  floating-point arithmetic operations, and requires  $(nz_U + (k+2)(n_S + 1) - 1)$  floating-point and  $(nz_U + n_S + 1)$  integer storage space for ADPH distributions of order  $n_S$ . Observe that when  $(I - P_{S,S})$  is upper-triangular,  $nz_U$  can be anywhere between  $n_S$  and  $n_S(n_S + 1)/2$ .

Next is an example of an acyclic DPH renewal process.

*Example 3. The discrete Erlang process (DEP).* Discrete Erlang process (DEP) is a DPH renewal process with  $l (= n_S)$  phases and the interrenewal distribution in Figure 2. It is represented by the initial probability vector  $\tau_S = (1, 0, \dots, 0)$  and the order  $(l + 1)$  DTMC

$$P_{DEP} = \left( \begin{array}{cccc|c} 1-a & a & & & \\ & 1-a & a & & \\ & & 1-a & \ddots & \\ & & & \ddots & a \\ & & & & 1-a & a \\ \hline & & & & & 0 & 1 \end{array} \right), \quad (7)$$



**Fig. 2.** The DEP interrenewal distribution

whose  $(l \times l)$  principal submatrix is acyclic; DEP reduces to the geometric distribution with parameter  $a$  when  $l = 1$ . In the general case, it can be conceived as the convolution of  $l$  independent geometric distributions with parameter  $a$ .

Using the acyclic  $(n_S \times n_S)$  principal submatrix of  $P_{DEP}$  in (7), we obtain the upper-bidiagonal matrix

$$I - P_{S,S} = \begin{pmatrix} a & -a & & & \\ & a & -a & & \\ & & a & \ddots & \\ & & & \ddots & -a \\ & & & & a \end{pmatrix}.$$

Then the moments of the DEP interrenewal distribution can be computed by following Algorithm 1 after its first step.

## 5 Numerical Results

In this section, we present results that illustrate the computation of moments of DPH distributions. We have implemented the proposed method in C using double-precision IEEE floating-point arithmetic [3], performed and timed all experiments under Cygwin 1.5.15-1 on a Pentium IV 3.4 GHz processor and a 1 GB main memory running Windows XP. In each problem, the degree of coupling,  $\|p_S\|_\infty$ , and the average degree of coupling,  $\|p_S\|_1/n_S$  (see (1)), are reported so as to indicate the inherent difficulty associated with computing the moments accurately [4].

We have experimented with two groups of problems. The first group consists of order 10, 20, 30, and 40 versions of MMBP (Example 2) and order  $10^1$ ,  $10^2$ ,  $10^3$ ,  $10^4$  versions of DEP (Example 3), and the particular initial distributions. The second group consists of order  $10^1$ ,  $10^2$ ,  $10^3$ ,  $10^4$  principal submatrices of the test suite of twelve DTMCs [6] that have been used in [5] and uniform initial distributions. In all problems, we have computed the first ten moments (whenever possible) in double precision floating-point arithmetic, but report results regarding only the first, second, fifth, and tenth moments in order to save space. The computation of binomial coefficients does not pose problems since



the largest entry in the tenth row of Pascal's triangle, 10 choose 5, is equal to, 252. Note that  $10^{308}$  is about the size of the largest positive number in double precision floating-point arithmetic, and any attempt to compute moment vectors with larger values than this will be futile. When we detect overflow in any of the moment vectors, we indicate it with n/a in the tables and do not compute further moment vectors.

All times are reported in seconds of CPU time. In the tables,  $T_s$  denotes the time to generate in core the absorbing DTMC in row-wise sparse format, to transform the rows corresponding to  $\mathcal{S}$  of the sparse DTMC so that elements of those rows are ordered according to column indices when the matrix is input from a file, and to determine the number of nonzeros in  $P_{\mathcal{S},\mathcal{S}}$ ;  $T_{LU}$  denotes the time to perform symbolic LU analysis on  $(I - P_{\mathcal{S},\mathcal{S}})$  and its sparse factorization (step 1 of Algorithm 1) as discussed in section 2; and  $T_m$  denotes the time to compute the first ten moments using Algorithm 1. Hence, the sum of  $T_{LU}$  and  $T_m$  is the time it takes Algorithm 1 to execute. In order to improve the confidence in the timing results, we have timed each experiment three times and have reported the average. In the tables,  $nz_{\mathcal{S}}$  denotes the number of nonzeros in  $(I - P_{\mathcal{S},\mathcal{S}})$ , which can be different than that in  $P_{\mathcal{S},\mathcal{S}}$  due to zero diagonal elements. We have reported  $nz_{\mathcal{S}}$  and  $nz_{LU}$  so that the amount of fill-in resulting from the LU factorization can be observed.

In  $P_{MMBP}$  and  $P_{DEP}$ ,  $p_{\mathcal{S}}$  has a single nonzero element such that  $\|p_{\mathcal{S}}\|_{\infty} = \lambda \sum_0$  and  $\|p_{\mathcal{S}}\|_{\infty} = a$  (see (6) and (7)), respectively. Hence,  $\|p_{\mathcal{S}}\|_1 = \|p_{\mathcal{S}}\|_{\infty}$  and for these two examples we only report  $\|p_{\mathcal{S}}\|_{\infty}$ . Furthermore, since  $(I - P_{\mathcal{S},\mathcal{S}})$  has symmetric nonzero structure in both examples and does not yield any fill-in during the factorization, it also turns out to be the case that  $nz_{\mathcal{S}} = nz_{LU}$ . The CPU timings  $(T_s, T_{LU}, T_m)$  in order 10, 20, 30, 40 versions of MMBP are all zero. The CPU timings in order  $10^1$ ,  $10^2$ ,  $10^3$  versions of DEP are all zero and the CPU timings in order  $10^4$  version of DEP are  $(T_s, T_{LU}, T_m) = (0, 3, 0)$ . Therefore, these timings do not appear in Tables 1 and 2.

Observe that it is more the average degree of coupling than the degree of coupling which determines the value of the first moment in MMBP with  $q = 1.0$  and DEP, and hence their other moments. There is a high degree of variability in the MMBP interrenewal distribution even for small orders and that moments exceeding  $10^{100}$  appear for order 30 or larger. On the other hand, the DEP interrenewal distribution approaches the deterministic distribution as the order increases. For  $a = 1.0$ ,  $P_{\mathcal{S},\mathcal{S}}$  has only an upper-diagonal resulting in immediate determinism. Recall that it takes only 3 seconds to find the first ten moments of the DEP interrenewal distribution of order  $10^4$  and that this time is spent in the first step of Algorithm 1.

Regarding the twelve test matrices used in the set of experiments reported in Table 5, we remark that they arise in six applications, namely 2D, pushout, mutex, ncd, qn, telecom, and that (easy, hard, medium), (mutex, mutex\_alt1, mutex\_alt2), and (ncd, ncd\_alt1, ncd\_alt2) are respectively three different versions of the pushout, mutex, and ncd problems. Furthermore, all test matrices except those of pushout are uniformized continuous-time MCs. With this second

**Table 1.** Moments of the MMBP interrenewal distribution

$a$	$q$	$\lambda$	$(n, nz)$	$n_S$	$(nz_S, nz_{LU})$	$\ p_S\ _\infty$	$E[H^1]$	$E[H^2]$	$E[H^5]$	$E[H^{10}]$
2.0	1.0	0.1	(11,30)	10	(28,28)	2.0e-4	5.1e4	5.3e9	4.3e25	4.6e53
		0.5				9.8e-4	1.0e4	2.1e8	1.4e22	5.0e46
		1.0	(11,29)			2.0e-3	5.1e3	5.3e7	4.6e20	5.4e43
	2.0	0.1	(11,21)			2.0e-4	1.0e4	2.1e8	1.3e22	4.6e46
		0.5				9.8e-4	2.0e3	8.4e6	4.3e18	4.7e39
		1.0	(11,20)			2.0e-3	1.0e3	2.1e6	1.3e17	4.6e36
	1.0	0.1	(21,60)	20	(58,58)	1.9e-7	1.0e8	2.2e16	1.5e42	5.9e86
		0.5				9.5e-7	2.1e7	8.8e14	4.9e38	6.1e79
		1.0	(21,59)			1.9e-6	1.0e7	2.2e14	1.6e37	6.1e76
	2.0	0.1	(21,41)			1.9e-7	1.0e7	2.2e14	1.5e37	5.8e76
		0.5				9.5e-7	2.1e6	8.8e12	4.9e33	6.0e69
		1.0	(21,40)			1.9e-6	1.0e6	2.2e12	1.5e32	5.8e66
	1.0	0.1	(31,90)	30	(88,88)	1.9e-10	1.6e11	5.2e22	1.3e58	4.3e118
		0.5				9.3e-10	3.2e10	2.1e21	4.2e54	4.4e111
		1.0	(31,89)			1.9e-9	1.6e10	5.2e20	1.3e53	4.4e108
	2.0	0.1	(31,61)			1.9e-10	1.1e10	2.3e20	1.7e52	7.4e106
		0.5				9.3e-10	2.1e9	9.2e18	5.5e48	7.6e99
		1.0	(31,60)			1.9e-9	1.1e9	2.3e18	1.7e47	7.4e96
	1.0	0.1	(41,120)	40	(118,118)	1.8e-13	2.2e14	9.7e28	6.2e73	9.6e149
		0.5				9.1e-13	4.4e13	3.9e27	2.0e70	9.9e142
		1.0	(41,119)			1.8e-12	2.2e13	9.7e26	6.2e68	9.7e139
	2.0	0.1	(41,81)			1.8e-13	1.1e13	2.4e26	1.9e67	9.4e136
		0.5				9.1e-13	2.2e12	9.7e24	6.2e63	9.6e129
		1.0	(41,80)			1.8e-12	1.1e12	2.4e24	1.9e62	9.4e126

**Table 2.** Moments of the DEP interrenewal distribution

$a$	$(n, nz)$	$n_S$	$(nz_S, nz_{LU})$	$\ p_S\ _\infty$	$E[H^1]$	$E[H^2]$	$E[H^5]$	$E[H^{10}]$
0.1	(11,21)	$10^1$	(19,19)	1.0e-1	1.0e2	1.1e4	2.2e10	2.6e21
	0.5			5.0e-1	2.0e1	4.2e2	5.2e6	8.7e13
	1.0			1.0e0	1.0e1	1.0e2	1.0e5	1.0e10
0.1	(101,201)	$10^2$	(199,199)	1.0e-1	1.0e3	1.0e6	1.1e15	1.5e30
	0.5			5.0e-1	2.0e2	4.0e4	3.4e11	1.3e23
	1.0			1.0e0	1.0e2	1.0e4	1.0e10	1.0e20
0.1	(1001,2001)	$10^3$	(1999,1999)	1.0e-1	1.0e4	1.0e8	1.0e20	1.0e40
	0.5			5.0e-1	2.0e3	4.0e6	3.2e16	1.0e33
	1.0			1.0e0	1.0e3	1.0e6	1.0e15	1.0e30
0.1	(10001,20001)	$10^4$	(19999,19999)	1.0e-1	1.0e5	1.0e10	1.0e25	1.0e50
	0.5			5.0e-1	2.0e4	4.0e8	3.2e21	1.0e43
	1.0			1.0e0	1.0e4	1.0e8	1.0e20	1.0e40

group of problems, we aim at having an understanding of the kind of moment values that arise in DPH distributions and assessing the performance of Algorithm 1 in computing these moments.

**Table 3.** Moments of DPH interrenewal distributions arising in twelve test matrices

Problem ( $n, nz$ )	$n_S$	( $nz_S, nz_{LU}$ )	$\ ps\ _\infty$	$\frac{\ ps\ _1}{n_S}$	( $T_s, T_{LU}, T_m$ )	$E[H^1]$	$E[H^2]$	$E[H^5]$	$E[H^{10}]$
2D	$10^1$	(25,42)	8.9e-1	2.7e-1	(0,0,0)	3.1e0	1.4e1	2.8e3	7.2e7
(16641,	$10^2$	(352,1746)	9.0e-1	8.9e-2		8.0e0	9.3e1	3.4e5	8.5e11
66049)	$10^3$	(3843,60340)	9.2e-1	2.8e-2		2.4e1	8.8e2	9.5e7	6.1e16
	$10^4$	(39588,1662849)	9.4e-1	5.8e-3	(0,5,1)	3.7e2	2.6e5	2.2e14	5.0e29
easy	$10^1$	(43,50)	9.8e-1	3.9e-1	(0,0,0)	1.9e0	4.5e0	1.0e2	4.5e4
(20301,	$10^2$	(596,1723)	9.8e-1	1.3e-1		5.5e0	4.2e1	4.3e4	1.0e10
140504)	$10^3$	(6648,57706)	9.9e-1	4.4e-2		1.6e1	3.6e2	9.9e6	6.5e14
	$10^4$	(68880,1866042)	9.8e-1	1.4e-2	(1,6,1)	4.9e1	3.5e3	3.1e9	6.3e19
hard	$10^1$	(43,50)	9.0e-2	3.6e-2	(0,0,0)	5.0e1	5.5e3	5.6e10	9.1e23
(20301,	$10^2$	(596,1723)	9.4e-2	1.2e-2		6.0e2	9.5e5	3.1e16	3.6e35
140504)	$10^3$	(6648,57706)	9.8e-2	4.0e-3		5.9e3	1.0e8	4.2e21	7.1e45
	$10^4$	(68880,1866042)	9.1e-2	1.3e-3	(1,6,1)	5.6e4	9.5e9	4.0e26	6.7e55
medium	$10^1$	(43,50)	5.0e-2	2.0e-2	(0,0,0)	1.9e3	7.9e6	4.0e18	4.2e39
(20301,	$10^2$	(596,1723)	6.9e-2	6.9e-3		6.9e12	9.6e25	2.0e66	9.9e134
140504)	$10^3$	(6648,57706)	9.0e-2	2.3e-3		3.0e41	1.8e83	3.1e209	n/a
	$10^4$	(68880,1866042)	5.4e-2	7.0e-4	(1,6,2)	1.1e134	2.4e268	n/a	n/a
mutex	$10^1$	(28,100)	2.2e-2	1.9e-2	(1,0,0)	1.0e2	2.2e4	1.9e12	9.7e26
(39203,	$10^2$	(464,9398)	2.1e-2	1.6e-2		7.2e2	1.1e6	3.6e16	3.6e35
563491)	$10^3$	(7296,833658)	2.1e-2	1.2e-2	(1,7,0)	2.8e4	1.7e9	2.7e24	1.9e51
	$10^4$	(109492,66540178)	1.5e-2	7.2e-3	(2,4415,10)	3.3e7	2.3e15	5.3e39	7.1e81
mutex_alt1	$10^1$	(28,100)	2.2e-5	1.9e-5	(1,0,0)	2.1e5	8.9e10	5.0e28	6.4e59
(39203,	$10^2$	(464,9398)	2.1e-5	1.6e-5		2.7e9	1.5e19	1.7e49	7.7e100
563491)	$10^3$	(7296,833658)	2.1e-5	1.2e-5	(1,7,0)	1.9e16	7.5e32	3.3e83	2.7e169
	$10^4$	(109492,66540178)	1.5e-5	7.2e-6	(2,4385,10)	2.6e25	1.4e51	1.5e129	5.9e260
mutex_alt2	$10^1$	(28,100)	2.2e-8	1.9e-8	(1,0,0)	2.1e8	8.9e16	5.1e43	6.5e89
(39203,	$10^2$	(464,9398)	2.1e-8	1.6e-8		2.7e15	1.5e31	1.8e79	8.1e160
563491)	$10^3$	(7296,833658)	2.1e-8	1.2e-8	(1,7,0)	1.9e28	7.5e56	3.3e143	2.7e289
	$10^4$	(109492,66540178)	1.5e-8	7.2e-9	(2,4526,47)	2.6e43	1.4e87	1.5e219	n/a
ncd	$10^1$	(34,56)	1.7e-2	2.0e-3	(0,0,0)	4.7e3	6.5e7	1.7e21	1.4e45
(23426,	$10^2$	(508,2662)	9.1e-2	8.8e-3		7.9e3	2.1e8	2.6e22	2.2e47
156026)	$10^3$	(6040,122706)	3.5e-1	1.7e-2		1.6e4	7.4e8	3.1e23	8.1e48
	$10^4$	(65444,5615910)	9.9e-1	2.7e-2	(1,33,1)	4.7e4	6.0e9	3.5e25	3.4e52
ncd_alt1	$10^1$	(34,56)	1.7e-2	2.0e-3	(0,0,0)	3.7e3	4.1e7	5.1e20	9.1e43
(23426,	$10^2$	(508,2662)	9.1e-2	8.8e-3		7.1e3	1.6e8	7.9e21	7.3e45
156026)	$10^3$	(6040,122706)	3.5e-1	1.7e-2		1.5e4	6.8e8	2.1e23	2.2e48
	$10^4$	(65444,5615910)	9.9e-1	2.7e-2	(1,33,1)	4.6e4	5.9e9	3.2e25	2.4e52
ncd_alt2	$10^1$	(34,56)	1.7e-2	1.9e-3	(0,0,0)	3.6e8	5.8e17	1.8e46	1.7e95
(23426,	$10^2$	(508,2662)	9.0e-2	8.8e-3		2.1e17	1.5e35	4.7e89	9.9e181
156026)	$10^3$	(6040,122706)	3.5e-1	1.7e-2		1.0e31	3.6e62	1.2e158	n/a
	$10^4$	(65444,5615910)	9.9e-1	2.7e-2	(1,33,4)	3.4e53	4.4e107	6.8e270	n/a
qn	$10^1$	(26,50)	4.4e-1	2.1e-1	(1,0,0)	9.1e0	2.1e2	2.2e7	1.8e17
(104625,	$10^2$	(388,3837)	6.2e-1	1.6e-1		3.9e1	5.2e3	1.1e11	6.1e24
593115)	$10^3$	(4688,285709)	6.2e-1	1.1e-1	(1,1,0)	4.2e2	7.0e5	2.6e16	3.4e35
	$10^4$	(53224,17064913)	6.2e-1	5.6e-2	(1,325,3)	4.3e4	6.0e9	1.2e26	6.0e54
telecom	$10^1$	(34,48)	4.2e-3	1.1e-3	(0,0,0)	8.3e3	1.6e8	8.8e21	2.3e46
(20491,	$10^2$	(442,1420)	7.4e-3	5.3e-4		3.1e7	2.0e15	4.0e39	4.3e81
101041)	$10^3$	(4814,43098)	1.7e-2	3.2e-4		1.6e21	5.4e42	1.5e108	5.5e218
	$10^4$	(49234,610098)	8.4e-2	2.3e-4	(1,3,1)	4.3e192	n/a	n/a	n/a

Six of the DTMCs (`ncd`, `ncd_alt1`, `ncd_alt2`, `mutex`, `mutex_alt1`, `mutex_alt2`) have symmetric nonzero structure; however, all DPH distributions arising from the twelve test matrices yield some fill-in. For the `mutex`, `mutex_alt1`, `mutex_alt2`, and `qn` test matrices,  $nz_{LU}$  for the order  $10^4$  DPH distribution is larger than  $10^7$ , thereby making it the most difficult to solve. Otherwise, the first ten moments of all DPH distributions of order  $10^3$  and smaller can be computed in 8 seconds. If the `mutex`, `mutex_alt1`, `mutex_alt2`, and `qn` problems are excluded, the first ten moments of DPH distributions of order  $10^4$  can be computed in 38 seconds. The time spent by Algorithm 1 after its first step in all of these problems is within 4 seconds. Hence, it is again the LU factorization which dominates the computation, and it becomes excessive in the `mutex`, `mutex_alt1`, and `mutex_alt2` problems with order  $10^4$  (note that the discrepancy in the  $T_{LU}$  timings in this case is less than 5% and is immaterial). But then again,  $10^4$  is a relatively large order to consider in practice. Furthermore, when overflow is encountered in computing a particular moment, we notice that the computation time for that particular moment in  $T_m$  increases considerably. See, for instance, `mutex_alt2` with order  $10^4$ . This situation can be attributed to the exceptions generated during the forward and backward solutions with the  $L$  and  $U$  factors in the body of the outer for-loop of Algorithm 1 when there is overflow.

As in the first set of experiments, we again observe that the smaller the average degree of coupling, the larger the moments become and the more difficult it is to obtain accurate results. However, through the particular DPH distributions arising in the twelve test matrices, we also understand that it is not only the average degree of coupling but also the values of the nonzeros in the  $L$  and  $U$  factors (and therefore the values in  $(I - P_{S,S})$ ) that determine how the moments grow. See, for instance, the value of the first moment in `telecom` with order  $10^4$ , which has an average degree of coupling  $2.3 \times 10^{-4}$ .

## 6 Conclusion

In this paper, we have presented a recently introduced method to compute higher moments of DPH distributions in row-wise sparse format and have conducted numerical experiments on a test suite of DTMCs with it. The method is based on factorizing a nonsingular matrix as large as the order of the DPH distribution using a form of GTH with delayed row updates and solving as many linear systems with different right-hand sides as the number of moments desired. It has been shown that the factorization is available explicitly up to a symmetric permutation when the DPH distribution is acyclic, and therefore need not be computed. Timing results have been given to indicate the efficiency of the proposed method. First moments much larger than the reciprocal of the average degree of coupling have been reported, especially when the order of the DPH distribution becomes large. When the order is  $10^4$  or larger, there are cases which show that it will be useful to consider a fill reducing ordering of the coefficient matrix since the amount of fill-in during factorization may become excessive.

## References

1. A. Bobbio, A. Horváth, M. Scarpa, and M. Telek. Acyclic discrete phase type distributions: properties and a parameter estimation algorithm. *Performance Evaluation*, 54(1):1–32, Sep 2003.
2. A. Bobbio, A. Horváth, and M. Telek. The scale factor: a new degree of freedom in phase-type approximation. *Performance Evaluation*, 56(1–4):121–144, Mar 2004.
3. T. Dayar. *Software for computing moments of phase type distributions*, 2005. Available from <http://www.cs.bilkent.edu.tr/~tugrul/software.html>.
4. T. Dayar and N. Akar. Computing moments of first passage times to a subset of states in Markov chains. *SIAM Journal on Matrix Analysis and Applications*, to appear. Also as Technical Report BU-CE-0414, Department of Computer Engineering, Bilkent University, Ankara, Turkey, Dec 2004. Available from <http://www.cs.bilkent.edu.tr/tech-reports/2004/BU-CE-0414.pdf>.
5. T. Dayar and W. J. Stewart. Comparison of partitioning techniques for two-level iterative methods on large, sparse Markov chains. Technical Report BU-CEIS-9805, Department of Computer Engineering and Information Science, Bilkent University, Ankara, Turkey, Apr 1998. Available from <http://www.cs.bilkent.edu.tr/tech-reports/1998/BU-CEIS-9805.ps.gz>.
6. T. Dayar and W. J. Stewart. *A test suite of Markov chains*, 2005. Available from <http://www.cs.bilkent.edu.tr/~tugrul/matrices.html>.
7. R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Publishing Company, Reading, Massachusetts, U.S.A., 1989.
8. W. K. Grassmann, M. I. Taksar, and D. P. Heyman. Regenerative analysis and steady state distributions for Markov chains. *Operations Research*, 33(5):1107–1116, Sep–Oct 1985.
9. B. R. Haverkort. *Performance of Computer Communication Systems: A Model-based Approach*. John Wiley & Sons, Chichester, England, 1998.
10. G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. SIAM Press, Philadelphia, Pennsylvania, U.S.A., 1999.
11. S. Nakazawa, K. Kawahara, S. Yamaguchi, and Y. Oie. Performance comparison with layer 3 switches in case of flow- and topology-driven connection setup. In *IEEE Global Telecommunications Conference, Rio de Janeiro, Brazil*, volume 1a of *GLOBECOM'99*, pages 79–86. IEEE Press, New York, U.S.A., Dec 1999.
12. M. F. Neuts. Probability distributions of phase type. In *Liber amicorum Prof. Emeritus H. Florin*, pages 173–206. Department of Mathematics, University of Louvain, Belgium, 1975.
13. M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Johns Hopkins University Press, Baltimore, Maryland, U.S.A., 1981.
14. S. Robert and J.-Y. Le Boudec. New models for pseudo self-similar traffic. *Performance Evaluation*, 30(1–2):57–68, Jul 1997.
15. W. J. Stewart. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton, New Jersey, U.S.A., 1994.
16. M. Telek and A. Heindl. Matching moments for acyclic discrete and continuous phase-type distributions of second order. *International Journal of Simulation: Systems, Science & Technology*, 3(3–4):47–57, Dec 2002.